

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2003, 13:02:56 ; Search time 29 Seconds
(without alignments)
1191.372 Million cell updates/sec

Title: US-10-027-000-2

Sequence: 1 MADIDVEALIKRLTAEKVD.....DSVALRGKFTVGETYWSGV 833

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Swissprot_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2052	46.7	845	1 BGLS_KLUMA	P07337 kluyveromyc
2	1445.5	32.9	818	1 BGLS_AGRU	P27034 agrobacteri
3	1145	26.1	754	1 BGLB_CLOTM	P14002 clostridium
4	722	16.4	880	1 BGL2_SACFI	P22507 saccharomyc
5	706	16.1	876	1 BGL1_SACFI	P22506 saccharomyc
6	646.5	14.7	765	1 BGLX_SALTY	P56078 salmonella
7	635.5	14.5	765	1 BGLX_ECOLI	P33363 escherichia
8	593	13.5	825	1 BGLS_HANAN	P46835 hansenula a
9	581	13.2	860	1 BGL1_ASFAC	P48825 aspergillus
10	451.5	10.3	830	1 BGLS_BUPI	P15084 butyrivibri
11	425	9.7	947	1 BGLS_RUMAL	P15885 rumiococcu
12	219.5	5.0	642	1 YRBD_BACSU	P40406 bacillus su
13	164.5	3.7	192	1 BGLS_SCHCO	P29091 schizophyll
14	151.5	3.5	330	1 NAGZ_VIBCH	P96157 vibrio chol
15	147.5	3.4	329	1 NAGZ_VIBBU	P96157 vibrio chol
16	147	3.3	343	1 NAGZ_YERPE	P81573 yersinia pe
17	144	3.3	341	1 NAGZ_ECOLI	P56067 escherichia
18	140	3.2	341	1 NAGZ_PASAU	P56067 escherichia
19	139	3.2	351	1 NAGZ_ECOLI	O96016 pasteurella
20	136	3.1	341	1 NAGZ_SALTY	O87216 salmonella
21	134	3.1	341	1 NAGZ_SALTY	O87216 salmonella
22	133.5	3.0	332	1 NAGZ_PSEAE	O912K0 pseudomonas
23	130	3.0	568	1 DCPY_ZYMO	P06672 pseudomonas m
24	129	2.9	351	1 NAGZ_HAEIN	P44955 haemophilus
25	128.5	2.9	598	1 HEXA_ALTSO	P48823 alteromonas
26	125	2.8	1058	1 CARB_FUSUN	O87486 fusobacteri
27	123.5	2.8	380	1 PEL_PSEMA	O51915 pseudomonas
28	121.5	2.8	634	1 DHSA_ARATH	O82633 arabidopsis
29	121.5	2.8	2647	1 FLNA_HUMAN	P21333 homo sapien
30	121	2.8	661	1 DHSA_DROME	O94523 drosophila
31	120.5	2.7	664	1 DHSA_HUMAN	P31040 homo sapien
32	119.5	2.7	1070	1 AGU_CANTS	P23064 candida tsu
33	119	2.7	1953	1 BIGA_SALTY	P25927 salmonella

34	119	2.7	2774	1 MAPA_RAT	P34926 rattus norv
35	118.5	2.7	599	1 LAC2_PHACU	O02075 thanatephor
36	118.5	2.7	1064	1 CARB_LACIA	O9cfv2 lactococcus
37	117.5	2.7	1178	1 RPOC_CLOPE	O93r87 clostridium
38	116.5	2.7	1064	1 CARB_LACIC	O32771 lactococcus
39	116	2.6	914	1 BPA_BACSU	P39793 bacillus su
40	116	2.6	1057	1 CARB_STAM	O99ur5 staphylococ
41	115.5	2.6	1058	1 CARB_LACPL	P77866 lactobacill
42	115	2.6	826	1 VILI_CHICK	P02640 gallus gall
43	115	2.6	1057	1 CARB_STAM	P58940 straphylococ
44	113.5	2.6	646	1 DHSA_CAEEL	O09508 ctenorhadi
45	113.5	2.6	1171	1 TR12_STRCO	O9rkb9 streptomyc

ALIGNMENTS

RESULT 1
BGLS_KLUMA STANDARD: PRT: 845 AA.
ID BGLS_KLUMA
AC P07337
DT 01-APR-1988 (rel. 07, Last Created)
DT 01-APR-1988 (rel. 07, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Beta-glucosidase precursor (EC 3.2.1.21) (gentliobase) (cellobiase)
DE (Beta-D-glucoside glucohydrolase).
OS Kluyveromyces marxianus (Yeast) (Kluyveromyces fragilis)
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_Taxid=4911;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12424;
RX MEDLINE=88210533; PubMed=2835179;
RA Raynal A., Gerbaud C., Francinques M.C., Guerinneau M.;
RT Sequence and transcription of the beta-glucosidase gene of
RT Kluyveromyces fragilis cloned in Saccharomyces cerevisiae.";
RL Curr. Genet. 12:175-184(1987).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
CC glucose residues with release of beta-D-glucose.
CC -1- PATHWAY: Cellulose degradation.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X05918; CAA29353.1; .
CC PIR: A29148; GLVX.
DR InterPro: IPR002772; GH_3C.
DR InterPro: IPR001764; GH_3M.
DR Pfam: PF00933; Glyco_hydro_3_1.
DR Pfam: PF01915; Glyco_hydro_3_C_1.
DR PRINTS: PR00133; GLYDRLASE3.
DR PROSITE: PS00775; GLYCOSYL_HYDROL_F3; 1.
KW Hydrolyase; Glycosidase; Cellulose degradation; glycoprotein; signal.
KW SIGNAL
FT CHAIN 1 ? 845
FT ACT_SITE 225 225
FT CARBOHYD 66 66
FT CARBOHYD 304 304
FT CARBOHYD 438 438
FT CARBOHYD 621 621
FT SEQUENCE 845 AA; 93916 MW; 78DAC1550D592B5 CRC64;
Query Match 46.7%; Score 2052; DB 1; Length 845;
Best Local Similarity 47.2%; Pred. No. 5.9e-125;
Matches 403; Conservative 152; Mismatches 269; Indels 30; Gaps 11;

OY		1	MIDIDEAIIKKTITLAEKVDLIAGDFWHTLKPKHGVPSLFETDGPNGVGRKTFENCVP	60
Db		1	MSRFDEQILSELINODEKISLSADVDPFHTKKIERGLIPAVVSGPBGINGTKFFDDVP	60
OY		61	AACFPGCGTSLGSFFNOCTLLLEAGCKMMKGFAIKSAHVILGPTINNORSPLDGREGPESTGE	120
Db		61	SGCFPGNOTGLASFDDRLDETLAGTKMLAKRSIAKNAVAVLIGPTTNORQLDGGRGEPSE	120
OY		121	DPLFLAGLAALIRGIQSFVGVAOTIKHFICANDODERNRMVOSITYTERALREIVALPOIA	180
Db		121	DPLIAMATSSIVYKMGQGEGIAATAVKHFECNDLEDQRSSNSIVSERALREIYLEPRRLA	180
OY		181	VRSOGAFATVTAANGINVCSCSENKRYLDGMKRKMWDGLMSMTYTSTTAAYAAGL	240
Db		181	VRIANPVCMITAIKNKYNGHCQSOKLLIDLIRDMKWGDMLMSDMPETYTTAAAINNGL	240
OY		241	DLEMPRPFRFGELTLKFANYSNGKPFITH--IDORAELYOFVK-----KCAASGYTENGPS	294
Db		241	DIEFPFGTRMRRIALVSHLSNRREDITTEDVDDRKQYLKAIRKYVDNLKETGTIVENGPE	300
OY		295	TYYNNTPETFAALLIRKYNENGVILLKNEENVLP.LSKKKTTLLVGBNQAQATYHGGSAALR	354
Db		301	SISNNKTESDILREIADSIYVLTKNNNYLTYSKERQRYHYIGPNAAKAKTSSSGGSASMN	360
OY		355	AYYAAPPEGLSKEOLETPSPSYVGAITYVP-PILEGOLTPDGA-----GNMRVFENE	407
Db		361	STYYVSPTGYIYNKLGEKVEDTVYGASHKSIGSLAESLLDAAPRADENAAGLAKRYSN	420
OY		408	PGETPRKHIDELF---FTKTDMLVDYVHPKRAADT---WYADEGSTYTADEDCYTELG	460
Db		421	P--VEERSDEPPFYVTKYVNSNHLPDFKHKEYDKRNKPYFEVILTIOQYIPOEDGDYIFES	478
OY		461	LYYCGTARAYVDDQ.LVDMATRKQYPGDAFFGSAARETGRIINLYKNTYFKLIEFGSAPT	520
Db		479	LQVYSSGLELYDELIDLQKHNOERGSPFCFAGTKERKTKLLTKKQYVNWVRVYEGSGPT	538
OY		521	YTLKCDTIYPVGGSLRVGGCVIQDOAQEIKESVALAKEHOVIICGLNMIDMETEGADRA	580
Db		539	SGLVGEF---GAGGQQAQVYKAIADDEETIRNAELAHHDAVAILJLGNEMETELEGDHE	595
OY		581	SKKLPGLVDLIADYAAPNPVTVMVMOGTPEEMPWLDPATPAVIOAWYGGNETGNSIADY	640
Db		596	NNDDLKRNENLEVRAALKANPMTVIVNOSGVEPPEPWLEEANALVOAWYGENELGNALDY	655
OY		641	VGDVNPSPGKSLSPPKLOONPAFLNPRFAGTLYGEDEVVYGRTRYEPADNVAPFPG	700
Db		656	LYGDVYVPMGKSLSSHPFKLQDNPAFLNFKTEFGVVYGEDIPFGYRYELEQKFAAFPEG	715
OY		701	HGLSTYTFAFSNLSYSHKDKLSVLSVKNKTY--SVPGAQVAAQLYVYRPLQAKINRPVKEL	759
Db		716	YGLSYTFPELIDISPDKYTDKIDISVDYKNTMGOKFAGSEVYQYIFYFSALN-SKYSRKYVEL	774
OY		760	KGFANVELOPBETAIVITEDEKVAAYAFDEBRDQMCVEKGDIYEVYSSSAKDGVALAR	819
Db		775	KGFEEVYHLEPEBEKTYVINELELKDAISYFENBELGKMHVEAGEYLVSVGTSS--DDILSY	831
OY		820	GKFTYGETYMWSGV	833
Db		832	KEFYVERKULYKGL	845
RESULT 2				
BGLS_AGRUU				
ID	BGLS_AGRUU	STANDARD:	PRT:	818 AA.
AC	PZ7034:			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	01-DEC-1992 (Rel. 24, last annotation update)			
DE	Beta-glucosidase (EC 3.2.1.21) (Gentibiobase) (celllobiobase) (Beta-D-			
CHG-1.	glucoside glucocytrolase).			
GN				
OS	Agrobacterium tumefaciens.			

Query Match	Best Local Similarity	Matches 334; Conservative 134; Mismatches 310; Indels 47; Gaps 19;
6	VEAIIKKTLEKVDYLLAGIDFWHFKALPKHGVSILRTDQPNQVR-CTKRFNCPACF 64	32.9%; Score 1445.5; DB 1; Length 818;
2	IDDLIDKLTLEQVSLSGADFWITTALEKRGVPIKATDGPNGARGGSLVGVKSCF 61	40.5%; Pred. No. 8.8e-86;
65	PCGSLSTFNQTLLEAGKMKREATAKSAHVILGPTINNQRSPILGKGFESIGEDPFL 124	
62	PVAIMLGATWDPLELIERAGVALGCAKSKASGYLLAPTVNIHRSGLGRNFECSDEPAL 121	
125	AGLGAALIRIGISQGVATIKHFLCNDQEDRRMNQVSIYTERLRLREYALPDIAYRDS 184	
122	TACCAVAVINGVSGVAATIKHFVANSSETERQTPSSDDERLTREYLPFEBAVKA 181	
185	QPGAFMTAYINGVSCSENPKYLDGMLRKEMGDGLIMSDWYTSSTTEAVVAGLDEM 244	
182	GKVAWMSYKNTLNGTYISENFWMLLTLYLREEMGPDGVVMSDMFESHSTAEIINNGLDEM 241	
245	PGPRFGEETLKFNVSNGKPRPIHYIIDORAREVLOFVKKCAASGVTENGPEITVN--NTPF 302	
242	PGPRWDREKELVAUVREGKVALETVRASARRILLRLERV--GAFEKAPDLAEHALDLPF 298	
303	TAALIRKNGNGGYILLNENNVLPISAKK--KTLIVPNAKQATYHGGSAALRAYAVT 360	
299	DRALIROLGAGAVLLKND--GVPLAKSSFDQIADVIPNANASRWMGSGSARIAAHYVS 357	
361	PFQGLSKOLETPBS--YTVGAYTT--VPPIIGBQCLTPDPAQRMKRVFNERPCTPNRH 416	
358	PLBEGIRALNSANSLRHAVGCNNNRLLIDVSGE--MTVEYFKG--RGFESRPVHNETVE 412	
417	IDELFF--TKTGMHLVDYHNPKAADTWADMGSTYATDEDCYTELGLVVGCPAKAYVD 473	

Db 413 KGEFFWDLPSGDLADLAF-----SARMTATFVPOETGHEHIGMTNAGIARLFVVDG 463

QY 474 QLVVNDATKOVPPDAPFGSRTRETRINLVKNTYKKEIEGSAFTYLLKDDTYIPGHS 533

Db 464 ELVVDYDGMWKKEENFEFGANSEORAVTLGAARRRVAYEY-EAPKASLDGINIC----- 518

QY 534 SLRVGGCKVDDOAEIEKSVALKAEHDOYITICAGLNADWETEGADRSKRLPGVDOLIA 593

Db 519 ALRFGVEKPLGD-AGTAEAVETARKSDIYLLVLRGENDTBDLDPKRLPGROBELIE 577

QY 594 DVAANENTVVMQGTPEEMPLDAPVIAQWYGNSTGSIADVFGDYNPSGKSL 653

Db 578 AVAETNENNVVVLGTGPIEMWIGKRVAVLQWMPGDELGNALADVLFGDVEPAGRLPQ 637

QY 654 SFPRKLDNNAFLN-----FTEGRLYGEDYVGVRYEEFADKDNVPEFGHLSYTFEA 709

Db 638 TEPKRLDNSALITDDSIYFGDGHVRYAGTIVGRHHDTRIELEPFGELGTYTRFT 697

QY 710 FSNLSVY-----HKDKLSVSLSVKNTGVSVAQVADLYVFLQAKINRVEKELGFANY 765

Db 698 WGAPOLSGTEMGADG-LTVYVDVTNIGDRAGSDVQVLYVHS-PNAVREPRFKELRAFAKL 755

QY 766 ELQGETKAVTIEOEKYVAAYDEERDQMCVEKDEYVSDS 810

Db 756 KLAFGATGAVLKIAPRDL-AYFDVEAGRPRADAGKVELIYAASA 799

RESULT 3

Db BGLB_CLOTM STANDARD; PRT; 754 AA.

AC P14002;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Thermolabile beta-glucosidase B (EC 3.2.1.21) (gentiobiase)

DE (cellobiase) (beta-D-glucoside glucohydrolase).

GN BGLB.

OS Clostridium thermocellum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium

OX NCBI_TaxID=1515;

RN [1]

RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.

RC STRAIN=ATCC 27405 / DSM 1237;

RX MEDLINE=89364694; PubMed=2505054;

RA Graebnitz F., Ruecknagel K.P., Seiss M., Staudenbauer W.L.;

RT "Nucleotide sequence of the Clostridium thermocellum bgib gene encoding thermostable beta-glucosidase B: homology to fungal beta-glucosidases.";

RT Mol. Gen. Genet. 217:70-76(1989).

CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-glucose residues with release of beta-D-glucose.

CC -1- PATHWAY: Cellulose degradation.

CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: X15644; CA333665.1; -

DR PIR: S04381; S04381.

DR InterPro: IPR002772; GH_3C.

DR InterPro: IPR001764; GH_3N.

DR Pfam: PF00933; Glyco_hydro_3; 1.

DR Pfam: PF01915; Glyco_hydro_3_C; 1.

DR PRINTS: PR00133; GLHYDLASE3.

DR PROSITE: PS00775; GLYCOSYL-HYDROL_F3; 1.

KM Hydrolase; Glycosidase; Cellulose degradation.

FT ACT_SITE 231 231 BY SIMILARITY.

SO SEQUENCE 754 AA; 84100 MW; ABCEB38BFC9FD3A8 CRC64;

Query Match 26.1%; Score 1145; DB 1; Length 754;

Best Local Similarity 31.9%; Pred. No. 2,1e-66;

Matches 270; Conservative 132; Mismatches 249; Indels 196; Gaps 15;

QY 4 IDVEAILKRLTLAEKVDLLAGIDFHWTKALPKHGVPSLRTFDGPNVGRTK-----FF 56

Db 3 VDIKTIKQMTLEEKALGCSGIDFHWTKRVERLTGIPSIIMTDGPHGLKREDEAEIADIN 62

QY 57 NCVPACPPCGSLGISTNQTLLAEAGKMKKEALAKSAHLYIPTIMORSPGGRGFE 116

Db 63 NSVPATCFPSAAGLACSDREIVERVGAALGECQAEVNSITLGGANIKSPGLGRNFE 122

QY 117 SGEPPFLAGLAALINGIOSTGVQATIKHPLCNDQEDRRMYOSITVERALREIYALP 176

Db 123 YPEDEPILSSELAASHIKINGVOSGAGCLKHPAANNQEHRRRTVOTIYDEKTIKEIYPAS 182

QY 177 FOIAVRDSOPGFMTAVNGINGVSCSENPKYIDGKLKREKMGDGLMSDWTGYSTTAV 236

Db 183 FENAVYKAPRVVVMCAVYNKNGEYSENRYLTLEVLKNEHMHDFVSDMGAVNDRVSGL 242

QY 237 VAGLDLEMPGPPRFGEITLKFNVSNKGPFIHYIDRAREVLOFVKCAASVTENGPRETT 296

Db 243 DAGLDLEMTSGITDDKILVEAVKSGKLENTLNAREVILKIYIM-----ALEKKENA 297

QY 297 VNNTEPETAALRKRVNGEIVLLKNENNVLPISKKRKTLLVGNPAQATYHGGGSAALRAY 356

Db 298 QYEQDAHHRHLAQAAESVNLKNEEDVLPFKKSGTIALIGAFVKPKPRYGSGSSHI--- 354

QY 357 YAVTPEDGLSKOLETPPSITVGAITYTPPILEGOCITPGAGMWRVFNPRGTPNQH 416

Db 355 -----TPTR-- 358

QY 417 IDELFETKTDMLVDYHHPKADTWYADMEGTYTADDECTYELGLVCGTAKAYVDOLY 476

Db 359 LDDI-----YELK-----KAGAD----- 372

QY 477 VDNATKQVPGDAFFGSATRETRINLVKNTYKKEIEGSAFTYLLKDDTYIPGHSILR 536

Db 373 -----KVLVYSEGR-----LENDGI----- 389

QY 537 VGGCKVIDOAEIEKSVALKAEHDOYITICAGLNADWETEGADRSKRLPGVDOLIAVYA 596

Db 390 -----DEELINEAKKKAASSDVAVPAAGLPDESESGFRTHTSIPEONRLIEAVA 441

QY 597 AANPTVVMQGTPEEMPLDAPVIAQWYGNSTGSIADVFGDYNPSGK 650

Db 442 EVGNSILVYVLLNSPYEMWIDKRVLEAYLGQALGGRWMCYSKSIY-----GK 494

QY 651 LSLSPKRLQDNPAFLNFTREAGRTLYGEDYVGVRYEYFADKDNVFPFGHLSYTFAF 710

Db 495 LAFTFPVKLSHNSPYLNFGEEDRVYKGLFEGVRYTDTKLEPLFPGHLSYTFKFX 554

QY 711 SNLSVSHKD-----GKLSVSLSVKNTGVSVAQVADLYVFLQAKINRVEKELGFAYE 766

Db 555 SDLSVKKDDVSDNSIINVSVAKNVGMKKELYVLVADVSS-VRRPEKELGKEXFV 613

QY 767 LQGETKAVTIEOEKYVAAYDEERDQMCVEKGDY-----EVIYSDSAAKDGYA 817

Db 614 LNGEKEKTYVF-TLDRARFAYVTQIKDMHVESGEFLILIGRSSDIVLKESRVNSTVK 672

QY 818 LRKGFYV 824

Db 673 IRRKFTV 679

RESULT 4

Db BGL2_SACFI STANDARD; PRT; 880 AA.

ID BGL2_SACFI

DT 01-AUG-1991 (Rel. 19, Created)

DT 01-AUG-1991 (Rel. 19, Last sequence update)

[illegible]

QY	219	DOILINDDWCVSTTEAAVYAGDLMPCG-----	-PREFGETLKFVNSGKPFRIHV	27
Db	293	QGFVSDMAAQAQSGAVSAISGLDMSMPGELLGAGMTKSGYSGQNLTKAVIETVPIERLD		
QY	271	QBARVLL-----QEVKCAASG--VTENGPEPTTVNN-----TP	30	
Db	353	DMATRIALALATNSFPKRDRLPNFSSTTEYGEHFFVDKTSPPVKKVNHVDPNSDTE	41	
QY	302	ETFAALLRKVNGNGVILLKNNVLPDS--KKKTLI---VGNPAK-----OAT	34	
Db	413	DIAL---KVAESIVYLLKNKEKTLPISEPKVKRLKLLSGIAGAPDPKGYECSDOSCVDAL	46	
QY	345	YHGSGSAALRAY--YAAMPFDGLSQLETPSPSYTGAVTVPRILEGCLPLPDGAPGRW	40	
Db	470	EGGMSGSV-GPKQVTPPEISA-----	49	
QY	403	RVFNEPPTPNRQHIDELFTKTDMLNDVYHPRKAADTVADMEGTATYADECYELGLV	46	
Db	494	---NARKNNMOPDYITRESF-----DLNQSVTASDAHMSIVV	52	
QY	463	VCGTAKAYVDQDLVVDNATKQVPGDAFGSATREBTGAINLVKNQTYFKLETGSAPITY	52	
Db	529	SAVSEGY---LIITD-----GN-----	54	
QY	523	LKGDITVGHGSLRVGCGKVIDDQAEIKESVALAKEHDOVITICAGLNADMETEGADRASM	58	
Db	543	RGD-----KNNVTL-----WHNS-----	55	
QY	563	KLPGLVDOLLADVAANPNVYVM-QTGTPEEMPLDTPRAYIAWYG--GNETGNSIAD	63	
Db	556	-----DMLIAVAENCANFTVYVITSQVGVSEFADHPNTALVYMGAPLDRSGTAIAN	60	
QY	640	VVEGDNVSGKLSLSFPRKDODNNAFLNFRFAG---RTLXGDDVYGVGYEYEFACKDV	69	
Db	610	ILFGANPSGHLPFTVAASNDYIPITYVNPNEPEDNTLAEHDLVDVRYFEKNIEP	66	
QY	696	NEPFGHGLSYTTFEAFNSLSVSH-----	71	
Db	670	RYAFGYGLSYMEKYSNAKKSAAKVDELPRLYLAEYSYNTKEIINN-EDAFPPSNA	72	
QY	718	-----KDGK-----LSVLSV	72	
Db	730	RIIOEFLPYLDNSVNTLKDGNYEYDPDGSTEQRTPIOPGGGLGNDALMEVAKVYEDV	78	
QY	729	KNTGSVPAQAOVAFOLYVPKLOAKINRPYKELKGFAYKELORGETKATYEQEKYVAAYF	78	
Db	790	QNLGNSIDKEFPQLYLKHEDGKFTETPY-QLNKGEKVELSPGEKTYEFELLRDLDSVW-	84	
QY	789	DEERDQWCEKGDYEVIVSDSSAAKDQVAL	818	
Db	848	DTTQOSWIVESGTYEAL-----GVAV	869	
RESULT 5				
BGL1_SACFI STANDARD; PRT; 876 AA.				
AC	P22506;			
DT	01-AUG-1991 (Rel. 19, Created)			
DT	01-AUG-1991 (Rel. 19, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Beta-glucosidase 1 precursor (EC 3.2.1.21) (Gentloblase) (Cellloblase)			
GN	(Beta-D-glucoside glucohydrolase).			
OS	BGL1			
OC	Saccharomycopsis fibuligera (Yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycopsidaceae; Saccharomycopsis.			
OX	NCBI_Taxid=944;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=89133518; PubMed=3146949;			
RT	Mechinda M., Ohtsuki I., Fukui S., Yamashita I.;			
RT	"Nucleotide sequences of Saccharomycopsis fibuligera genes for			


```

QY 484 VPGCAFGSGSATREETGRINLVKGMTYKFKIEFGSAPYTLTKGDIIVPBGSLRGCKVY 543
Db 484 DPR----- 486
QY 544 DDQAEIEXSVALAEHQVYITCAGLNMDETREGADRSMTKPGVLDOLIAVAAANPTY 603
Db 487 SPQAMIDEAVQAAQADVAVAVVGEESQMAHEASSRNITITPOSORDITLTKTKPLV 546
QY 604 VVMQGTPEEPMWLD-ATPAVIOAWGCGENGSNIADVEFGDYVPSGKLSLSPFKRLDN 662
Db 547 LVLNNGRPLALVKEDQADALILETFMFACTEGCNMIADVLPEDYVNSGKLPISFPRVGOI 606
QY 663 PAFLNFTTEAGRTLYGE-DVYVGYRYRYEFKDKVNEFPFGHLSYTFPAFNSLSV---- 716
Db 607 PVYYS-HLNTGRPNVPEKPNKTYTS-RYFDEANGCLP-VYFGGLSYTFETVSDVLSSPTM 663
QY 717 HKDCKLSVSLSVKNTGSGVPGQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 776
Db 664 QRDRETVASVETVNTGREGATVYIOMLYODV-TASMSFPAVQMLGFEKTIYLPGRKTVS 722
QY 777 -----IEOEKRYVAVYFDEERDQMCVEKGDYEVYVS-DSSAAKDG 815
Db 723 FPIDE-----ALKFVNQQAKYDAECPGFNVFVIGVDSARVKQG 760

RESULT 7
ID BGLX_ECOLI STANDARD; PRT; 765 AA.
AC P3363;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Periplasmic beta-glucosidase precursor (EC 3.2.1.21) (Gentloblase)
DE (Cellloblase) (beta-D-glucoside glucosylolase).
DE BGLX OR B2132.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_Taxid=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / W3110;
RA Yang M., Luch S., Goddard A., Reilly D., Henzel W., Bass S.;
RL Submitted (SPP-1994) to the EMBL/GenBank/DBJ databases.
RP [2]
RC SEQUENCE FROM N.A.
RC STRAIN=K12 / BHB2600;
RA Richterlich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,
RL Church G.M.;
RT Submitted (CCT-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; Pubmed=9278503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
CC glucose residues with release of beta-D-glucose.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).

```


Query Match 13.5%; Score 593; DB 1; Length 825;
 Best Local Similarity 24.5%; Pred. No. 1,1e-30;
 Matches 209; Conservative 111; Mismatches 262; Indels 272; Gaps 26;

9 ILKLTIAEVNLLAGIDP-----FTKALPKHGVSLFTDGPNGVGTKEFPNGVPA 62
 63 LVDDMSIAEKVNLTTGGASGPGSGNTGSPRLNLSICVODGPIPLSVRAADLT----- 117
 63 CFCPGTSLGTFNQTLLLEAGKMKKEAIAKSAHVILGPTIN--MOSPLGREGFESIGED 121
 118 VPCGMMASSFNKQLIYDAVAIAKGFRRKGDALIGFYGMGKAAAGRWEGHDP 177
 122 PRLAGLAALIRIGIOSTGVQATIKHFLCNDQEDRRM-----MVOS 162
 178 PYLEGVTAIYQITIGIOSGVSTAKHLIGNEHEFRFAKKDKHAKIDPOMENTSSLS 237
 163 IYERAIAREIYALPFOJAVNDSQPGAFMTATYNGINGVSCSENKYIDGMIRKMGWDGLI 222
 238 EIDDRAMHEIYLPFAVAVNG--VSSIMCSYKLNKNGSHACONSULLNLYLKELEFGQGV 296
 223 MSDMYGTSTTEAVVAGLDLEMPGPRFRGETLKFVNSGKPFIVHIDORAREVLQFVK 282
 297 MTDMGALYSGIDANANGLDMDMPCENQYFGNLTAVLNGITLPQRLDMATRILSAL-- 354
 283 CAASGV--TENG-----ETVYNTPEPTALLRKVNGE 314
 355 -IYSGVHNDGPNYMAQOTLEEGHEXFKOEGDIVLNKHVVRSDINRAVALRS--AVBG 412
 315 IYLLKNNVLPKSKKKTLI-----VGPNAK--QATYHGGSSALRAVYA----- 358
 413 VVLLKNEHELPLRGREKVKRISILGGAADDGSGTSCIRGCGSGAIGTGYSGAGTESY 472
 359 -VTPREGLSKOLETPPSYVGAATYVPIIGEOCLIPDGAFGMRKRVENPEPTENRQHI 417
 473 FYTPADGI-----GARAQOEKISYEFIG----- 495
 418 DELFFKTMHLVDYHPRADRTYADMEGTATADECTYELGLVYVCGIAKAYVDQVY 477
 496 -----DSM-----NOAAMASALYA 510
 478 DNATKQVPGDAFFGATRETRGRINLVKNTYKFKIEFSAPTYTLKGPITYPGHSLRV 537
 511 DAIEVA-----NSVAGEIGDVG--GN-----YGDINLNL----- 540
 538 GGCKVTDQAEIEKSVAAKEDQYITICAGLNADMETEGADRAKMLPGVLDQILIADYAA 597
 541 -----MHNAY-----PLIKNIS 553
 598 ANPNVYVVGOTGPPEM--PWLD--ATPRAVIAQWYNGNETGNSIADYVGDYVPSKLSLS 654
 554 INNNTIYVTSQOQIDLEFFIDNEVNTAVITYSSIGQFVYLAFLVGLDENPESKLPFT 613
 655 FPKRLQDN--PAFLNRTGAGRTLYGEDYVGYRYEFAFDKVNPFPGSLSTYTFAFENL 713
 614 IAKDVNDYIPVIEKYVDPPVVKFTESIYDVRFDKTKKVRVYRFFGSLSTNSFSLDI 673
 714 SVSHDKLSVLSKATNGSVPGAOVAQLYKPLQAAKINRVELKFAVAVELQPGFTK 773
 674 EQ-----TLQPSSENAEPANYSFY-----QYKSNMPPSE-- 706
 774 AVTIEOEKRYAAV 787
 707 -YTVPEGFKELANY 719

RESULT 9
 BGLI ASPAC STANDARD; PRT; 860 AA.
 AC P48825;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 13-JUN-2002 (Rel. 41, Last annotation update)

DE Beta-glucosidase 1 precursor (EC 3.2.1.21) (Gentiobiase) (Cellbiase)
 DE (Beta-D-glucoside glucosylhydrolase).
 OS Aspergillus aculeatus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus;
 NCBI_TaxID=5053;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97082984; PubMed=8964516;
 RA Kawaguchi T., Enoki T., Tsurumaki S., Sumitani J., Ueda M.,
 RA Ooi T., Arai M.;
 RT "Cloning and sequencing of the cDNA encoding beta-glucosidase 1 from
 RT Aspergillus aculeatus";
 RL Gene 173:287-288(1996).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
 CC glucose residues with release of beta-D-glucose.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL; D64088; BAA10968.1; -
 DR InterPro; IPR002772; GH_3C.
 DR InterPro; IPR001764; GH_3N.
 DR Pfam; PF00933; Glyco_hydro_3; 1.
 DR Pfam; PF01915; Glyco_hydro_3_C; 1.
 DR PRINTS; PR00133; GLHYDRLASE3
 DR PROSITE; PS00775; GLYCOSYL_HYDROL_3; 1.
 KW Hydrolyase; Glycosidase; Cellulose degradation; glycoprotein; signal.
 FT SIGNAL 1 19
 FT CHAIN 20 860
 FT ACT_SITE 20 280
 FT CARBOHYD 61 61
 FT CARBOHYD 211 211
 FT CARBOHYD 252 252
 FT CARBOHYD 315 315
 FT CARBOHYD 322 322
 FT CARBOHYD 354 354
 FT CARBOHYD 387 387
 FT CARBOHYD 442 442
 FT CARBOHYD 523 523
 FT CARBOHYD 542 542
 FT CARBOHYD 564 564
 FT CARBOHYD 658 658
 FT CARBOHYD 668 668
 FT CARBOHYD 690 690
 FT CARBOHYD 712 712
 SQ SEQUENCE 860 AA; 93052 MW; 4B464778B00FC694 CRC64;

Query Match 13.2%; Score 581; DB 1; Length 860;
 Best Local Similarity 23.3%; Pred. No. 7e-30;
 Matches 225; Conservative 114; Mismatches 299; Indels 326; Gaps 28;

8 ALLKLTIAEVNLLAGIDP-----FTKALPKHGVSLFTDGPNGVGTKEFPNGVPA 60
 49 AIVSQMTLDERKVNLTGTG--WELEKCVGOTGGVPRNLNIGMCLDPSLPLRISDY----- 102
 61 AACFCGTSLSGTFNQTLLLEAGKMKKEAIAKSAHVILGPTIN--MOSPLGREGFESIGED 119
 103 NSAFPGAVYNAATWMDKNLAVLRQAMGQEPFSDKIDVQLGPAAGPLGRSPDGGRWEGFS 162
 120 EDPFLAGLAALIRIGIOSTGVQATIKHFLCNDQEDRRM-----VOSIYVE 166
 163 PDEALGVLFATRIKIGIODGAVATKHYILNDEHFRQVABAAGYGNISDTISSNVD 222
 167 RALREIYALPFOJAVNDSQPGAFMTATYNGINGVSCSENKYIDGMIRKMGWDGLMSDW 226

RX MEDLINE=90175009; PubMed=2106673;
 RA Ohmura K., Takano M., Shimizu S.;
 RT "DNA sequence of a beta-glucosidase from *Ruminococcus albus*.";
 RL Nucleic Acids Res. 18:671-671(1990).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
 CC glucose residues with release of beta-D-glucose.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X15415; CA3461.1; -
 DR PIR: S08243; S08243.
 DR InterPro: IPR002772; GH_3C.
 DR InterPro: IPR001764; GH_3N.
 DR Pfam: PF00933; Glyco_hydro_3; 1.
 DR Pfam: PF01915; Glyco_hydro_3_C; 1.
 DR PRINTS: PR00133; GLHYDRLASE3.
 DR PROSITE: PS00775; GLYCOSYL_HYDROL_F3; 1.
 DR Hydrolyase; Glycosidase; Cellulose degradation.
 FT ACT_SITE 696
 FT ACT_SITE 696 BY SIMILARITY.
 FT SEQUENCE 947 AA; 104277 MW; C43B8CDD9D60A115 CRC64;
 Query Match 9.7%; Score 425; DB 1; Length 947;
 Best Local Similarity 34.7%; Pred. No. 9, 6e-20;
 Matches 107; Conservative 50; Mismatches 105; Indels 46; Gaps 6;
 QY 31 KALPHGVPSLRFDPNGV---GTFKGVNPAACPCGCTSLGTFNQLLEAGKMMG 87
 DB 505 KHELEGLIPACCSGDSGSMRLDVGTRAFS-----LPNGLIAATFNKSLTELEPTIYG 558
 QY 88 KEATKSAHVILGPTLNNORSPGLGRFESIGEDPEPLAGAAALLRGIOSTVOATYH 147
 DB 559 LEMRANKYDCLLGGGMNTHHPLNGRFEFFSEDPFLTGMAALEGLHSGVEGTIKH 618
 QY 148 FLCDQEDRRMVOISYTERIALPQIAVRDSDPGAFATVANGINGVSCSNPKY 207
 DB 619 FCANQETNRHFIIDSVASERALEIKLGEIYAVRSKASVMTGYKNGMLTMSFPL 678
 QY 208 LDGMKREMGWDLIMSDWYGTST-----TEAVVAGIDLEMPGPPRFGTLEKF 257
 DB 679 NTMLRKQMGDFDTMTDMANINDKGCAPDKNNFAAMVRAQNDVYMCAD--GSGSD 735
 QY 258 NV---SNGKPFTHVIDQAREVLYQFKKCAASGVTEGNETTNNPTETALLRKVME 313
 DB 736 NVIALADGRLTRAEIQRASRNILSFM-----STHMAKRLGED 775
 QY 314 GIVLKEE 321
 DB 776 EAVEVINK 783
 RESULT 12
 YBBD_BACSU STANDARD; PRT; 642 AA.
 AC P40406;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical lipoprotein ybbd precursor (ORF1).
 GN YBBD.
 OS *Bacillus subtilis*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_Taxid=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BD99 / MS94;

RX MEDLINE=94281248; PubMed=8011666;
 RA Quirk P.G., Guffanti A.A., Clejan S., Cheng J., Krulwich T.A.;
 RT "Isolation of Tn917 insertion mutants of *Bacillus subtilis* that are
 RT resistant to the protonophore carbonyl cyanide
 RT m-chlorophenylhydrazine.";
 RL Biochim. Biophys. Acta 1166:27-34(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168.
 RX MEDLINE=97419514; PubMed=9274029;
 RA Liu H., Haga K., Yasumoto K., Ohashi Y., Yoshikawa H., Takahashi H.;
 RT "Sequence and analysis of a 31 kb segment of the *Bacillus subtilis*
 RT chromosome in the area of the *rnh* and *rnc* operons.";
 RL Microbiology 143:2763-2767(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bortner M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Bortner M., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Frits S.Y., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Katamata D., Katsuhara Y., Kikuchi-Bianhard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kunita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestel D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetle D., Porwollik S., Prescott A.M.,
 RA Pressac E., Puig P., Purnelle B., Rapoport G., Ray M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sedate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
 RA Seliguchi J., Sekowski A., Seror S.J., Seror P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccoti E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandemol M., Vannier F., Vassaretto A.,
 RA Vialti R., Wambuit R., Wedler E., Wedler H., Weitzenger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*.";
 RL Nature 390:249-256(1997).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (potential).
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: L19954; AAA64351.1; -
 DR EMBL: AB002150; BAA19499.1; -
 DR EMBL: J99104; CAB11942.1; -
 DR Subtilist; BG10832; Ybbd.
 DR InterPro: IPR002772; GH_3C.
 DR InterPro: IPR001764; GH_3N.
 DR Pfam: PF00933; Glyco_hydro_3; 1.
 DR Pfam: PF01915; Glyco_hydro_3_C; 1.
 DR PROSITE: PS00775; GLYCOSYL_HYDROL_F3; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Hypothetical protein; Hydrolase; Glycosidase; Membrane; Lipoprotein;
 KW Signal; Complete proteome.

FT SIGNAL 1 16 POTENTIAL.
 FT CHAIN 17 642 HYPOTHEICAL LIPOPROTEIN YBBD.
 FT LIPID 17 17 N-ACYL DIGLYCERIDE (POTENTIAL).
 FT ACT_SITE 318 318 BY SIMILARITY.
 SQ SEQUENCE 642 AA; 70580 MW; DCEA93142922P13F CRC64;

Query Match
 Best Local Similarity 21.5%; DB 1; Length 642;
 Matches 101; Conservative 75; Mismatches 126; Indels 167; Gaps 20;

QY 4 IDVEALLKTLTAEKVDLACIDF--W-----
 DB 33 IDANQIVNRSLEDEKIGLMDPFRMVGKGESSPOLATKMNDEVALYKKYQFGGIIIF 92
 QY 29 -----HTKLPKHGVPSLRFETDGPNGV-----RGKFFNGVPAACPG 67
 DB 93 AENKTKTQVLTLDIYQKSPK--IPLMISIDGEGIVRLGEGTN-----PFGN 141
 QY 68 TSLGTFNQTLEBAGKMGKAKSAHVILGPTINMRSP-----LGKGFESIGEDPF 123
 DB 142 MALGAASRLNAVQTSIIIGKELSGALGINDFSPVDINNPNPIYGVRSFSSNELTS 201
 QY 124 LAGIGAAALRGIOSTGVQVATTHFLCNDGEDRRMAYQ-----IYT--ERALREIYALP 176
 DB 202 RLGG--YTKKGLQKODIASALKHFFPGHGDTD---VDSHYGLPIYSHGGERLREVELY 254
 QY 177 FOIARNDGOGAFMTAY-----NGING---VSCSNPRTYLDGMLRKGMDG 220
 DB 255 FQKAI-DAGADWMTAVHVPAPDPTTKKSLDGLVATSKVYGLLQEMGFNG 313
 QY 221 LIMSWMYGT-----YSTEAVV---AGLDLMPGPPRPGETLKNVNSNGKPFTHVID 270
 DB 314 YIVTDALNMKAIAIDHFGOEAVVAVAKAGVIAL-----MPASVSLKE-----E 358
 QY 271 QRAVEVQFPKCKAASVTENGPEPTYNNPETAALL----- 307
 DB 359 QKFAVYIOLKAEVKNK--DIPQOQINNSVERTISLIKRGWYPANRSDSTKEIKAK 415
 QY 308 -----RKVNEGIVLLKNNVPLSKK--KTLIVGPNAKQ 342
 DB 416 KIYSGKHLEAKELAKAVTLKNEQHTLPFRPKKGRILIIYAPYEQ 464

RESULT 13
 BGLS SCHCO STANDARD; PRT; 192 AA.
 ID BGLS SCHCO
 AC P20931;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Beta-glucosidase (EC 3.2.1.21) (Gentlobiose) (Cellulobiose)
 DE (beta-D-glucoside glucosylhydrolase) (Fragment).
 OS Schizopyllum commune (Bracket fungus).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 OC Agaricales; Schizophyllaceae; Schizophyllum.
 OX NCBI_TaxID=5334;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=66295810; PubMed=3091028;
 RA Moraneil F., Barbier J.R., Dove M.J., Mackay R.M., Sellig V.L.,
 RA Yaguchi M., Willick G.E.;
 RT "A clone coding for Schizopyllum commune beta-glucosidase: homology
 RT with a yeast beta-glucosidase.";
 RL Biochem. Int. 12:905-912(1986).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
 CC glucose residues with release of beta-D-glucose.
 CC -1- PATHWAY: Cellulose degradation.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M27313; AAA33925.1; -
 DR PIR: A28571; A28571.
 DR InterPro: IPR002772; GH_3C.
 DR InterPro: IPR001764; GH_3N.
 DR Pfam: PF01915; Glyco_hydro_3_C; 1.
 DR PROSITE: PS00775; GLYCOSYL_HYDROL_F3; PARTIAL.
 KW Hydrolase; Glycosidase; Cellulose degradation; Glycoprotein.
 FT NON_TER 1
 SQ SEQUENCE 192 AA; 28410C1805E648A5 CRC64;

Query Match
 Best Local Similarity 3.7%; DB 1; Length 192;
 Matches 40; Conservative 19; Mismatches 52; Indels 5; Gaps 3;

QY 589 DDLIADVAANPNVYVMOT--GTPEEMWLDATPAVIAWYG--GNETGNSIADVFGDY 645
 DB 77 DALVQAVADANENTIIAANTVGAITTEAWIEHPNKAAYVSGLPQEGAGNSVADILYGAY 136
 QY 646 NPSGKLSLFPKRLDNPALFNRTEA--GRLYGEDVYVGYRYEFPDKVNPFE 699
 DB 137 NPSGRLPYIAKSADYPAQVLYESSAQVPIIDYSEGLVDYRHPDANGIEPRFE 192

RESULT 14
 MAG2_VIBCH STANDARD; PRT; 330 AA.
 ID MAG2_VIBCH
 AC Q9K037;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Beta-hexosaminidase (EC 3.2.1.52) (N-acetyl-beta-glucosaminidase)
 DE (Beta-N-acetylhexosaminidase).
 GN MAG2 OR VC0692.
 OS Vibrrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eissen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Dragoi I., Sellers P.,
 RA Ermolaeva M.D., Yamathayan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Uterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
 RT cholerae.";
 RL Nature 406:477-483(2000).
 CC -1- FUNCTION: Cleaves GlcNAc linked beta-1,4 to MURNAc tripeptides
 CC (by similarity).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal non-reducing N-
 CC acetyl-D-hexosamine residues in N-acetyl-beta-D-hexosaminides.
 CC -1- PATHWAY: CELL WALL SYNTHESIS; MUREIN TRIPEPTIDE RECYCLING PATHWAY.
 CC -1- SUBUNIT: MONOMER (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES. MAG2
 CC SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL: AE004155; AAF93857.1; -

DR TIGR: VC0692; -
 DR Interpro: IPR001764; GH_3N.
 DR Pfam: PF00933; Glyco_hydro_3; 1.
 DR PROSITE: PS00775; GLYCOSYL_HYDROL_F3; 1.
 DR Hydrolase: Glycosidase; Peptidoglycan synthesis; Cell division;
 KM Cell wall; Complete proteome.
 FT ACT_SITE 242 242
 SQ SEQUENCE 330 AA; 36465 MW; A737BD2C149D3A2 CRC64;

Query Match 3.5%; Score 151.5; DB 1; Length 330;
 Best Local Similarity 22.8%; Pred. No. 0.0099;
 Matches 75; Conservative 44; Mismatches 111; Indels 99; Gaps 12;

QY 8 AILKILAEKVDLLAGIDFWHTKALPKHGVPRLFTDGPNGVGRKFFNGVPACPGCG 67
 DB 44 ALNKAIQAQAKRPILIGVD-----QEGGRVQRFREG-----FSRIPPAQYVAR 86
 QY 68 TSLGSTNQLLEAGMMKKEAIKSAHYILGPTIN--QRSPLGGRGESIGEDPFLA 125
 DB 87 AENK-----VELAEGGWLMAAELIAHDIDLSFAPVLDKGPACAKINRAF---GEVQTV 139
 QY 126 GLGAALIRIGSTGVQATIKHF-----LCNDEDRRMVQSIYTERALREIYA 174
 DB 140 LKSSAFLRGKAVGATTKHPPGHGAVIADSHLETPYDERETIQDMAIFRAQIEAGV 199
 QY 175 LPFOIAVR-----DSQPGAFMTAVNGINGVSCSENPKYIDGLKREKMGDLINSDWYG 228
 DB 200 LDAMPAPHVVPYHYDAP-----ASGSSYWLKQVIRELGFKIGFSD--- 242
 QY 229 TYSTTEVAVAGLDEMPGPRFRGETLKFNVSKRPFIVIDORARE-----VLQTV 280
 DB 243 -----DLSMEGAAMVKGPPV-----ERSHQALVAGCDMLIGN 274
 QY 281 KKCAASGVTEGPEPTVNNPTETALLRK 309
 DB 275 KREAIVEVDNL---IMEVPAQAEALLKK 300

RESULT 15
 NAG2_VIBFU STANDARD; PRT: 329 AA.
 AC P96157;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Beta-hexosaminidase (EC 3.2.1.52) (N-acetyl-beta-glucosaminidase)
 GN (Beta-N-acetylhexosaminidase).
 OS NAG2 OR EXO II.
 OC Vibrion furiosus.
 OX Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrion.
 RN NCBI_TaxID=29494;
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-16.
 RC STRAIN=7225;
 RX MEDLINE=97125984; PubMed=8969206;
 RA Chitlariu E., Roseman S.,
 RT "Molecular cloning and characterization of a novel beta-N-acetyl-D-
 RT glucosaminidase from *Vibrio furnissii*.";
 RL J. Biol. Chem. 271:33433-33439(1996).
 RN [2]
 RP REVISIONS TO 70-80.

RA Chitlariu E., Roseman S.;
 RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Cleaves GlcNAc linked beta-1,4 to MurNAc tripeptides
 CC (By similarity). HYDROLYSES RAPIDLY P-NITROPHENYL-N-ACETYL-BETA-D-
 CC GLUCOSAMINIDE (PNP-BETA-GLCNAC) AND 4-METHYLBELLIFERYL-BETA-
 CC GLCNAC, AND SLIGHTLY ACTIVE ON P-NITROPHENYL-BETA-GALNAC. MAY PLAY
 CC A ROLE IN SIGNAL TRANSDUCTION BETWEEN HOST AND ORGANISM.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal non-reducing N-
 CC acetyl-D-hexosamine residues in N-acetyl-beta-D-hexosaminides.
 CC -1- ENZYME REGULATION: INHIBITED BY GLCNAC, 2-ACETAMIDO-1-N-(4-L-
 CC ASPARTYL)-2-DEOXY-BETA-D-GLUCOPYRANOSYLAMINE (GLCNAC-ASN) AND O-
 CC (2-ACETAMIDO-2-DEOXY-D-GLUCOPYRANOSYLIDENE)-AMINO-N-

CC PHENYLCARBAMATE (PNCMC).
 CC -1- PATHWAY: CELL WALL SYNTHESIS; MOREIN TRIPEPTIDE RECYCLING PATHWAY.
 CC -1- SUBUNIT: MONOMER (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- MISCELLANEOUS: MAXIMAL pH AND OPTIMAL TEMPERATURE OF THE ENZYME
 CC ARE 7.0 AND 45 DEGREES CELSIUS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES. NAG2
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: U52818; AAC44686.2; -
 DR Interpro: IPR001764; GH_3N.
 DR Pfam: PF00933; Glyco_hydro_3; 1.
 DR PROSITE: PS00775; GLYCOSYL_HYDROL_F3; 1.
 KM Hydrolase; Glycosidase; Peptidoglycan synthesis; Cell division;
 FT ACT_SITE 242 242
 SQ SEQUENCE 329 AA; 36181 MW; 440C6A9B18143C34 CRC64;

Query Match 3.4%; Score 147.5; DB 1; Length 329;
 Best Local Similarity 23.3%; Pred. No. 0.018;
 Matches 73; Conservative 43; Mismatches 114; Indels 83; Gaps 11;

QY 16 AKVULLAGIDFWHTKALPKHGVPRLFTDGPNGVGRKFFNGVPACPGCGTSLGSTFN 75
 DB 52 AKRPILIGVD-----QEGGRVQRFREG-----FSKIPADLVARSNGT--- 91
 QY 76 QTLLEAGMMKKEAIKSAHYILGPTIN--MORSPLGGRGESIGEDPFLAGGAALI 133
 DB 92 -OLAEDEGWLMAAELIAHDIDLSFAPVLDKGPDCRAIGNRAF---GDDYCVLYTSSAYM 147
 QY 134 RGIQSTGVQATIKHF-----LCNDEDRRMVQSIYTERALREIYAALPFOIAVR 182
 DB 148 RGMKSVGATTKHPPGHGAVIADSHLETPYDERDSIADMDITFRAQIEAGVILDMMPAH 207
 QY 183 -----DSQPGAFMTAVNGINGVSCSENPKYIDGLKREKMGDLINSDWYGTSTIEAV 236
 DB 208 VIYPHYDAP-----ASGSPYWLKQVIRELGFQIGFISD----- 242
 QY 237 VAGLDEMPGPRFRGETLKFNVSKRPFIVIDORAREVLPFKKCAASGVTEGPEPTT 296
 DB 243 -----DLSMEGAAMVKGPPRAQGS-----LDAGCDMYLMCKKRESAVAVLDOLPISV 290
 QY 297 VNNPTETALLRK 309
 DB 291 V---PQAQSLKKQ 300

Search completed: April 26, 2003, 13:08:55
 Job time : 36 secs